applied sciences

MDPI

Article

Analysis of the Composition of Ancient Glass and Its Identification Based on the Daen-LR, ARIMA-LSTM and MLR Combined Process

Zhi-Xing Li [1,†], Peng-Sen Lu [1,†], Guang-Yan Wang [1,*], Jia-Hui Li [1], Zhen-Hao Yang [1], Yun-Peng Ma [1] and Hong-Hai Wang [2,*]

[1] School of Information Engineering, Tianjin University of Commerce, Tianjin 300134, China; 17837858022@163.com (Z.-X.L.); 18177753918@163.com (P.-S.L.); jiahuilee1211@163.com (J.-H.L.); 120210545@stu.tjcu.edu.cn (Z.-H.Y.); mayunpeng@tjcu.edu.cn (Y.-P.M.)
[2] School of Chemical Engineering and Technology, National-Local Joint Engineering Laboratory for Energy Conservation in Chemical Process Integration and Resources Utilization, Hebei University of Technology, Tianjin 300130, China
* Correspondence: wanggy@tjcu.edu.cn (G.-Y.W.); ctstwhh@hebut.edu.cn (H.-H.W.); Tel.: +86-139-0211-3897 (G.-Y.W.)
† These authors contributed equally to this work.

Abstract: The glass relics are precious material evidence of the early trade and cultural exchange between the East and the West. To explore the cultural differences and trade development between early China and foreign countries, it is extremely important to classify glass cultural relics. Despite their similar appearances, Chinese glass contains more lead, while foreign glass contains more potassium. In view of this, this paper proposes a joint Daen-LR, ARIMA-LSTM, and MLR machine learning algorithm (JMLA) for the analysis and identification of the chemical composition of ancient glass. We separate the sampling points of ancient glass into two systems: lead-barium glass and high-potassium glass. Firstly, an improved logistic regression model based on a double adaptive elastic network (Daen-LR) is used to select variables with both Oracle and adaptive classification characteristics. Secondly, the ARIMA-LSTM model was used to establish the correlation curve of chemical composition before and after weathering and to predict the change in chemical composition with weathering. Thirdly, combining the data processed by the above two methods, a multiple linear regression model (MLR) is used to classify unknown glass products. It was shown that the sample obtained by this processing method has a very good fit. In comparison with other similar types of models like Decision Trees (DT), Random Forests (RF), Support Vector Machines (SVM), and Random Forests based on classification and regression trees (CART-RF), the classification accuracy of JMLA is 97.9% on the train set. The accuracy rate on the test set reached 97.6%. The results of the research demonstrate that JMLA can improve the accuracy of the glass type classification problem, greatly enhance the research efficiency of archaeological staff, and gain a more reliable result.

Keywords: Daen-LR; ARIMA-LSTM; MLR; machine learning; cultural heritage; ancient glass classification

check for updates

Citation: Li, Z.-X.; Lu, P.-S.; Wang, G.-Y.; Li, J.-H.; Yang, Z.-H.; Ma, Y.-P.; Wang, H.-H. Analysis of the Composition of Ancient Glass and Its Identification Based on the Daen-LR, ARIMA-LSTM and MLR Combined Process. *Appl. Sci.* **2023**, *13*, 6639. https://doi.org/10.3390/app13116639

Academic Editors: Giulia Festa and Claudia Scatigno

Received: 21 April 2023
Revised: 16 May 2023
Accepted: 29 May 2023
Published: 30 May 2023

1. Introduction

Machine learning (ML) algorithms are a set of mathematical models and statistical [1] methods that can be used in computer systems to learn and make predictions or decisions based on patterns in data. In the field of archaeology, there are many examples of machine learning algorithms applied in the direction of conservation and restoration, provenance research, and the management of cultural heritage. In 1798, the German scientist M.H. Klaproth conducted the first quantitative chemical study of Roman-era glass [2], improving the procedure for weight analysis and devising various procedures for the determination of non-metallic elements, accurately determining the composition of nearly 200 minerals and various industrial products. In 2003, Professor Fuxi Gan and his research team used the proton-excited X-fluorescence (PIXE) technique to quantify the chemical composition of a batch of ancient glass excavated in Yangzhou and Hubei, with the goal of studying the origin, system, and preparation process of ancient Chinese glass [3]. As more and more ancient silicate artifacts were unearthed, some scholars began to classify them based on their chemical composition. In 1992, Korean scientist Lee Chul applied his chemometric pattern recognition method to multivariate data to determine the classification of 94 ancient Korean glass pieces using neutron activation analysis and principal component analysis [4]. In 2010, El-Taher, an Egyptian scholar, used instrumental neutron activation analysis (INAA) and HPGe detector γ-spectroscopy to determine qualitatively and quantitatively for the first time a total of 16 elements in feldspar rock samples collected from Gabel El Dubb, Eastern Desert, Egypt, and to classify their rock samples [5]. In 2011, Thai scholar Won-in K. and his team used Raman spectrophotometry for the first time to characterize fragments of archaeological glass samples with the aim of obtaining information to identify glass samples for classification by laser scattering [6]. In 2019, Nadine Schibille and her team established a temporal model that serves as a tool for dating archaeological glass assemblages as well as a geographical model that allows for a clear classification of Levantine and Egyptian plant ash glasses [7]. However, it is worth noting that the application and extension of machine learning algorithms in the direction of cultural heritage (CH) component analysis and identification of categories are very limited [8].

In recent years, when studying the chemical composition of ancient glass objects, the classification of glass has been mainly determined by the weight ratio of oxides or by analyzing the mass fraction of compounds containing lead and potassium [9–13]. However, the percentage of lead and potassium compounds present varies depending on the region where the glass was produced and the degree of weathering, which would interfere with the classification of the glass. Thereby, this study is based on the data related to ancient glassware provided by the official website of the 2022 China Student Mathematical Modeling Competition [14]. The weathering of glass over thousands of years can cause significant changes in its internal chemical composition. As a result, determining the type of glass by the amount of content in a certain chemical composition is not reliable or scientific. Therefore, based on the double adaptive elastic net improved logistic regression model (Daen-LR), ARIMA-LSTM model, and multiple linear regression model (MLR) [15,16], optimizing and combining these three algorithms, we propose a processing method and process that is suitable for classifying complex data and can be used to predict the unknown classification of glass. This process method is used to analyze and model the data related to the chemical composition and classification information of a batch of ancient Chinese glass products, to find out the correlation between their chemical composition and the basis of their classification, and to use this relationship to predict the category of unknown glass. The accuracy of the algorithm model was judged by testing the presence of heteroskedasticity in the perturbation terms, testing for multicollinearity, and testing the fit of the experimental values through the model to the actual values [17].

With the continuous development of machine learning technology, a variety of machine learning models have been proposed and widely used in classification research. These include Logistic Regression (LR), Naive Bayes(NB), Decision Tree(DT), Support Vector Machine(SVM) and Random Forest(RF), Gradient Boosting Tree(GBT), and so on. However, traditional machine learning methods have some drawbacks in solving real-world problems, such as interference from external factors, failure to meet scientific standards, random results, and poor prediction accuracy. In order to solve these problems, it is necessary to combine sophisticated machine learning methods with more advanced methods. This paper presents a joint machine learning algorithm using Daen-LR, ARIMA-LSTM, and the MLR model (JMLA). We first use an improved logistic regression model based on double adaptive elastic networks (Daen-LR) to select variables that have both Oracle and adaptive classification properties. Secondly, we use the ARIMA-LSTM model to balance the linear and nonlinear trends in the time series data of the chemical content of glass artifacts before and after weathering. Finally, a multiple linear regression model (MLR) was used to classify the experimental samples. By testing the data set of the 2022 Chinese College Students Mathematical Contest in Modeling, this study proves the correctness of the proposed method.

The main contributions of this study include:

1. Successfully established a classification model of ancient glass products with high accuracy.
2. This study combines three different algorithms reasonably and effectively and integrates the advantages of different algorithms into the JMLA algorithm.
3. We made a comprehensive comparison of multiple test sets on multiple models, and the test results show that the algorithm given in this study is superior to other algorithms.
4. In the future, this algorithm model will also be able to support component analysis in many fields, such as water flow pollution, food safety, and environmental protection.

This paper consists of six parts: In the second part, the algorithm and principles of this paper are described in detail. In the third part, the preprocessing of the data and the preparation of the experiment are explained. In the fourth part, the experimental process and results are discussed. In the fifth part, the advantages and limitations of the JMLA model compared with other models are given. In the sixth part, the conclusion, influence, and future research suggestions are drawn.

2. Theory and Method

2.1. An Improved Logistic Regression Model with Double Adaptive Elastic Net

In this analysis of ancient glass artifacts, the relationship between glass weathering and its chemical composition was identified and statistically analyzed by using an improved logistic regression model based on a double adaptive elastic net, i.e., the Daen-Logistic regression (Daen-LR) model. Furthermore, we calculated the *p*-value of each correlation factor and counted whether each regression coefficient was significant at the 90% confidence level, extracted the strong correlation elements, and excluded the weak correlation elements. The characteristics of the model are as follows:

Logistic regression is an effective method to solve classification problems in which effective estimation of parameters and selection of variables are extremely important. The regularization method [18], which considers adding a penalty term to the optimized loss function to estimate parameters, can simultaneously solve the two key points of logistic regression. Elastic net [19] is one of the representatives of this method.

However, considering the inadequacy of the traditional logistic regression model for the estimation of parameters and the identification of important variables, it has two major shortcomings: First, the selected variables may not be consistent, i.e., they lack oracle properties [20]. Second, the specific effects of strongly correlated variables on the independent variables are not considered, i.e., adaptive categorical effects are missing [21,22].

To overcome the first deficiency of Elastic net, adaptive elastic net is established by combining Adaptive lasso [23] and Ridge to achieve consistency in selected variables. However, the Adaptive coefficient vector W1, which makes an adaptive elastic net with oracle properties, is not easy to set correctly. It is generally determined by the initial estimates of parameters and the constant δ.

To solve the second defect of Elastic net, Van et al. [24] proposed a Generalized ridge in which parameters are first divided into groups and then given different Ridge penalties for each group. The Generalized ridge has an adaptive grouping effect, and its Adaptive ridge also enjoys that effect. However, Generalized ridge does not have the function to select variables and is of limited application.

Based on existing solutions to the Elastic net deficiency, it follows that Adaptive lasso and Adaptive ridge have oracle properties and adaptive grouping effects, respectively, so they can be combined to avoid the two existing disadvantages. This combination of penalties can be called the double adaptive elastic net.

It is assumed that in the composition analysis of glass artifacts, there are m chemical composition influence factors, $X = (x_1, x_2, \cdots, x_m)$ is the characteristic variable of chemical composition content (i.e., independent variable), m is the number of variables, and the weathering status of the corresponding glass artifacts is set as y (i.e., dependent variable), where y represents the dichotomous variable of weathering or not (i.e., 0 means "unweathered" and 1 means "weathered"). To assess the magnitude of the probability of whether a particular glass artifact is weathered, it is necessary to calculate the predicted outcome of the model as the probability of occurrence of $y = 1$, which can then be expressed as $P = f(y = 1 \mid x_1, x_2, \cdots, x_m)$, i.e., the mathematical expression of the traditional logistic regression model is:

$$\mathrm{Logit}(P) = \ln\frac{P(y=1)}{1-P(y=1)} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m x_m \tag{1}$$

i.e.,

$$P(y=1) = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m x_m)}{1+\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m x_m)}$$

In the above equation, $(\beta_0, \beta_1, \beta_2, \cdots, \beta_m)$ are the regression coefficients to be determined. The great likelihood estimation method is used to find these coefficients:

$$P(y = 1 \mid \mathrm{X}) = 1 - \frac{1}{1+\exp(\beta_n + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m x_m)} = \pi \tag{2}$$

$$P(y = 0 \mid \mathrm{X}) = 1 - \pi \tag{3}$$

Combining the probability functions of y as:

$$P(y_i) = \pi^{y_i}(1-\pi)^{y_i}, y_i = 0, 1; i = 1, 2, \cdots, n \tag{4}$$

According to the Bernoulli distribution, the maximum likelihood function can be expressed as:

$$l(\beta; \mathrm{X}) = \prod_{i=1}^{n} P(y_i) = \prod_{i=1}^{n} \pi_i^{y_i}(1-\pi_i)^{1-y_i} \tag{5}$$

The log-likelihood function is expressed as:

$$\ln(l(\beta; X)) = \sum_{i=1}^{n}\{y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_m x_{im}) - \ln[1+\exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_m x_{im})]\} \tag{6}$$

Since the maximum likelihood function is a convex function, the point at which its first derivative equals zero is the point of maximum value. By calculating the first derivative of the undetermined coefficient $(\beta_0, \beta_1, \beta_2, \cdots, \beta_m)$ in Equation (6) and setting it equal to zero, all the parameters to be solved in the equation group can be solved.

Considering the shortcomings and deficiencies of the traditional logistic regression model, the oracle effect and adaptive classification effect are integrated into the traditional logistic model to create a double adaptive elastic net model, which makes the identification results of glass cultural relics more relevant and persuasive [25,26].

Theory 1. *Oracle property*

In the Logistic model, suppose the real parameters $\beta_0 = (\beta_{01}, \beta_{02}, \cdots, \beta_{0m})^T$, $\mathcal{A} = \left\{j \mid \beta_{0j} \neq 0\right\}$ $= \{1, 2, \cdots, m_0\}, m_0 < m$, *Fisher information matrix* $I(\beta_0) = \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix}$, *where* I_{11} *is a square matrix of order* m_0, $\phi\left(\mathbf{X}^T\boldsymbol{\beta}\right) = \ln\left(1 + \mathrm{e}^{\mathbf{X}^T\boldsymbol{\beta}}\right)$, *then the double adaptive elastic net logistic has oracle property according to the following conditions.*

1. $I(\beta_0)$ Is a positive definite matrix.
2. There exists an open set containing $\boldsymbol{\beta}_0$, such that for any $\boldsymbol{\beta} \in \Omega$ there exists a function $N(\cdot)$ satisfying:

$$\mid \phi'''\left(\boldsymbol{X}^T\boldsymbol{\beta}\right) \mid \leq \boldsymbol{N}(\boldsymbol{X}) < \infty, \tag{7}$$

and for any m-dimensional vector u, we have $E\left(N(\boldsymbol{X})\left(\boldsymbol{X}^T\boldsymbol{u}\right)^3\right) < \infty$;

3. $\lambda_1 = o(\sqrt{n})$, and there is a sequence $\{a_n\}$, such that:

$$a_n\left(\hat{\boldsymbol{\beta}}^* - \boldsymbol{\beta}_0\right) = O_p(1) \text{and} \lim_{n\to\infty}\frac{\lambda_1 a_n^\delta}{\sqrt{n}} = \infty;$$

4. $\lambda_2 = o(n)$ and $\lim_{n\to\infty}\frac{\lambda_2}{\sqrt{n}}\sqrt{\sum_{j=1}^{m_0}\beta_{0j}^2} = 0$.

When conditions 1–4 hold, double adaptive elastic net estimate $\hat{\boldsymbol{\beta}}$ has the following properties:

1. $\sqrt{n}\left(\hat{\boldsymbol{\beta}}_{\mathcal{A}} - \boldsymbol{\beta}_{\mathcal{A}}\right) \xrightarrow{D} N\left(0, \boldsymbol{I}_{11}^{-1}\right)$;
2. $\lim_{n\to\infty} P\left(\hat{\boldsymbol{\beta}}_{\mathcal{A}^c} = 0\right) = 1$.

Theory 2. *Adaptive Classification Effect*

Given the binary data $\{(\boldsymbol{X}_i, y_i)\}_{i=1}^n$, *where* $\boldsymbol{X}_i = (x_{i1}, x_{i2}, \cdots, x_{im})^T$ *and* $\forall j \in \{1, 2, \cdots, m\}$, $\sum_{i=1}^n x_{ij} = 0$, $\sum_{i=1}^n x_{ij}^2 = 1$, $y_i \in \{0, 1\}$. *Let* $\hat{\boldsymbol{\beta}}(\lambda_1, \lambda_2)$ *be the estimate of the model and assume that* $\hat{\boldsymbol{\beta}}_k(\lambda_1, \lambda_2)\hat{\boldsymbol{\beta}}_l(\lambda_1, \lambda_2) > 0$. *Define* $D_{\lambda_1,\lambda_2}(k, l) = \frac{1}{n}\left|w_{2k}\hat{\beta}_k(\lambda_1, \lambda_2) - w_{2l}\hat{\beta}_l(\lambda_1, \lambda_2)\right|$, *then:*

$$D_{\lambda_1,\lambda_2}(k, l) \le \frac{\sqrt{2(1-\rho_{kl})} + \frac{\lambda_1}{n}|w_{1k} - w_{1l}|}{2\lambda_2} \tag{8}$$

where

$$\rho_{kl} = \mathrm{corr}(x_k, x_l)$$

By combining the above two schemes to improve the logistic regression equation, $\boldsymbol{X}_i = (1, x_{i1}, x_{i2}, \cdots, x_{im})^T$, $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \cdots, \beta_m)^T$, $y_i \in \{0, 1\}$, $i = 1, 2, \cdots, n$, its estimated value of β is:

$$\hat{\boldsymbol{\beta}}_{\mathrm{Daen}} = \underset{\boldsymbol{\beta}}{\operatorname{argmin}}\left\{\ln(l(\beta; X)) + \lambda_1\sum_{j=1}^m w_{1j}\left|\beta_j\right| + \lambda_2\sum_{j=1}^m w_{2j}\beta_j^2\right\} \tag{9}$$

where

$$w_{1j} = \left|\hat{\beta}_j^*\right|^{-\delta}, w_{2j} > 0, \lambda_1, \lambda_2 > 0, \delta > 0$$

$$\hat{\boldsymbol{\beta}}^* = \underset{\beta}{\operatorname{argmin}}\left\{\ln(l(\beta; X)) + \lambda_2\sum_{j=1}^m w_{2j}\beta_j^2\right\}. \tag{10}$$

Since incorporating the correlation of variables into the regression model helps to improve the accuracy of parameter estimation and variable selection [27],

$$w_{2j} = \frac{\sum_{k=1,k\ne j}^m\left|\rho_{kj}\right|}{m-1} + \varepsilon_j,$$

where $\rho_{kj} = corr\left(x_k, x_j\right)$ is the correlation coefficient between variables x_k and x_j, $(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_m)^T$ is the vector that can make $w_{21}, w_{22}, \cdots, w_{2m}$ unequal to each other, and $\frac{\sum_{j=1}^m \varepsilon_j}{m} = 1, 0.95 \le \varepsilon_j \le 1.25$.

Equation(9) is equivalent to:

$$\begin{gathered}\hat{\boldsymbol{\beta}}_{\mathrm{Daen}} = \underset{\beta}{\operatorname{argmin}}\{\ln(l(\beta; X))\}, \\ \text{s.t.}\alpha\sum_{j=1}^m w_{1j}\left|\beta_j\right| + (1-\alpha)\sum_{j=1}^m w_{2j}\beta_j^2 \le t,\end{gathered} \tag{11}$$

where

$$\alpha = \frac{\lambda_1}{\lambda_1 + \lambda_2}, t > 0.$$

Using the coordinate gradient method and the Newton method to solve $\boldsymbol{\beta}$, Equation (9) can be rewritten as:

$$\hat{\boldsymbol{\beta}}_{\text{Daen}} = \underset{\boldsymbol{\beta}}{\text{argmin}}\left\{ I(\boldsymbol{\beta}) + \lambda_1 \sum_{j=1}^{m} w_{1j}\left|\beta_j\right| \right\}$$

where

$$I(\boldsymbol{\beta}) = \ln(l(\beta; X)) + \lambda_2 \sum_{j=1}^{m} w_{2j}\beta_j^2.$$

If $\boldsymbol{\beta}(t)$ is the solution of $\boldsymbol{\beta}$ at step t, then $I(\boldsymbol{\beta})$ can be approximated as:

$$I(\boldsymbol{\beta}) \approx I(\boldsymbol{\beta}(t)) + (\boldsymbol{\beta} - \boldsymbol{\beta}(t))^T \boldsymbol{g}(t) + \frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\beta}(t))^T \boldsymbol{h}(t)(\boldsymbol{\beta} - \boldsymbol{\beta}(t)) \tag{12}$$

where $\boldsymbol{g}(t)$ and $\boldsymbol{h}(t)$ are the gradients of $I(\boldsymbol{\beta})$ at $\boldsymbol{\beta} = \boldsymbol{\beta}(t)$, respectively, and the Hessian matrix, adding $\lambda_1 \sum_{j=1}^{m} w_{1j}\left|\beta_j\right|$ to the Equation (12) and making $\frac{\partial I(\beta)}{\partial \beta} = 0$ yields:

$$\boldsymbol{\beta}(t+1) = K\left(\boldsymbol{\beta}(t) - \boldsymbol{h}^{-1}(t)\boldsymbol{g}(t), \lambda_1 \boldsymbol{h}^{-1}(t)\boldsymbol{W}_1\right) \tag{13}$$

where

$$\boldsymbol{W}_1 = (0, w_{11}, w_{12}, \cdots, w_{1m})^T,$$

$$K(Q, W) = \begin{cases} Q - W, & 0 \leq W < Q \\ Q + W, & 0 \leq W < -Q \\ 0, & |Q| \leq W \end{cases} \tag{14}$$

Since $\lambda_1 \boldsymbol{h}^{-1}(t)\boldsymbol{W}_1$ may have some numbers less than 0, the parameters of some irrelevant variables cannot become 0. Thus, it can be directly rewritten as $\lambda_1 \boldsymbol{W}_1$. By the above inference, the solution process of $\hat{\boldsymbol{\beta}}_{\text{Daen}}$ can be derived: first generate an initial value of $\boldsymbol{\beta}$, then repeat the calculation of $\boldsymbol{g}(t)$, $\boldsymbol{h}(t)$, and $\boldsymbol{\beta}(t+1)$, until convergence [26].

2.2. Time Series Forecasting Model Based on ARIMA-LSTM

2.2.1. ARIMA(p,d,q) Model

By using the ARIMA(p,d,q) model, it is possible to analyze observations at past time points, depict the intrinsic link between them, and predict future values, which is achieved based on past time values and linear error equations [28–32]. The ARIMA model is usually denoted as ARIMA(p,d,q), where p is the number of autoregressive terms, q is the number of sliding average terms, and d is the number of differences needed to make it a smooth series [33]. The correlogram, autocorrelation function (ACF), and partial autocorrelation function (PACF) of the time series provide information about the lags [34]. If the time series is found to be smooth, the model can be used for estimation and forecasting. However, if it is not smooth, in order to apply ARIMA, it must be transformed to be smooth by differencing. After identification, an ARIMA model is estimated for a specific smooth time series. The simple ARIMA model is estimated based on the number of effective coefficients, the Bayesian information criterion (BIC) and the Akaike information criterion (AIC), and the adjusted R^2 [35]. After estimation, the selected ARIMA model needs to be diagnosed to check if the residuals are white noise. If the residuals are not white noise, the model must be re-estimated, and Q-tests and normality tests can be used to diagnose the residuals [36]. Typically, the ARIMA model is as follows:

$$y_t' = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i}' + \varepsilon_t + \sum_{i=1}^{q} \beta_i \varepsilon_{t-i} \tag{15}$$

$$y_t' = \Delta^d y_t = (1 - L)^d y_t \tag{16}$$

$$\left(1 - \sum_{i=1}^{p} \alpha_i L^i\right)(1 - L)^d y_t = \alpha_0 + \left(1 + \sum_{i=1}^{q} \beta_i L^i\right)\varepsilon_t \tag{17}$$

For the study, the more the number of chemical content parameters, the better the model fit, but this will be at the cost of increasing the model complexity, so the model selection should seek the best balance between the model complexity and the ability of the model to explain the data. According to the Bayesian information criterion, when the BIC is smallest, the optimal solution between the fit effect and complexity can be found [37]:

$$\mathrm{BIC} = \ln(\mathrm{T})(\mathrm{n}) - 2\ln(\mathrm{M}) \tag{18}$$

T: number of samples;
n: number of unknown parameters, n = p + q + 1;
M: maximum likelihood number of the model.
The maximum likelihood estimation process for the ARIMA(p,d,q) model is [38]:

$$Y_t = c + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \cdots + \Phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \tag{19}$$

$\Phi_1, \Phi_2 \cdots \Phi_p$ respectively represent the autolinear correlation coefficients between $Y_{t-1} \cdots Y_{t-p}$ and Y_t. By introducing the $p-1$ term in the middle, the direct relationship between $Y_{t-1} \cdots Y_{t-p}$ and Y_t can be separated, and this relationship is linear. $\Phi_1, \Phi_2 \cdots \Phi_p$ is the value to measure the size of this influence, which is the so-called PACF($\theta_1 \cdots \theta_q$ is the same meaning as $\Phi_1, \Phi_2 \cdots \Phi_p$). ε_t is the perturbed term. Where $\varepsilon_t \sim \mathrm{iidN}(0, \sigma^2)$, The vector of total parameters is

$$\vec{\Theta} = \left(c, \Phi_1, \Phi_2, \ldots, \Phi_p, \theta_1, \theta_2, \ldots, \theta_q, \sigma^2\right)$$

The estimation of the likelihood function for the autoregressive process is conditioned on the initial value of y, and the estimation of the likelihood function for the moving average process is conditioned on the initial value of ε. Then ARIMA(p,d,q) is conditioned on d as the difference order and the initial values of y and ε.

Assume the initial values $\vec{y_0} = (y_0, y_{-1}, \ldots, y_{-p+1})'$ and $\vec{\varepsilon_0} = (\varepsilon_0, \varepsilon_{-1}, \ldots, \varepsilon_{-q+1})'$ is known, then according to $\{y_1, y_2, \cdots, y_r\}$, it can be iterate to this equation:

$$\varepsilon_t = y_t - c - \Phi_1 y_{t-1} - \Phi_2 y_{t-2} - \cdots - \Phi_p y_{t-p} - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \tag{20}$$

The sequence $\{\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_T\}$ for $t = 1, 2, \cdots, T$ can be obtained, then the conditional likelihood function is:

$$L(\theta) = \ln f_{Y_T, Y_{T-1} \ldots, Y_1 | \vec{Y}_0, \vec{t}_0}\left(y_T, y_{T-1}, \ldots, y_1 \mid \vec{Y}_0, \vec{\varepsilon}_0, \vec{\theta}\right) = -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln\left(\sigma^2\right) - \sum_{t=1}^{T} \frac{\varepsilon_t^2}{2\sigma^2} \tag{21}$$

2.2.2. LSTM Model

ARIMA(p,d,q) model can well deal with the linear part of the chemical composition content in the time series, but it has certain limitations because the obtained residual series results have nonlinear characteristics, and the process of the content of some chemical components changing with the degree of weathering is a nonlinear process. This requires a deep learning model to solve the nonlinear trend of chemical composition changes [39]. The LSTM (Long Short-Term Memory) model is a deep learning model that is very good at solving nonlinear data. Its nonlinear gate unit can adjust the information flowing into and out of memory tuples at each time point so as to better fit the trend of nonlinear data changing over time.

LSTM is a special type of recurrent neural network (RNN) that performs very well with long sequences of data, mainly solving gradient disappearance, gradient explosion, and overfitting problems when training long sequences [40–43]. RNN is an artificial neural network that operates on time-series data and can use back-propagation algorithms to learn and adapt to the relationship between inputs and outputs. In contrast to standard RNN, LSTM has an input gate, a forgetting gate, and an output gate that control the way information flows through the network. These gates of the LSTM allow it to store past information and update the current state appropriately, thus providing a significant advantage when dealing with long sequences of data. The basic structure is shown in Figure 1 [44,45].

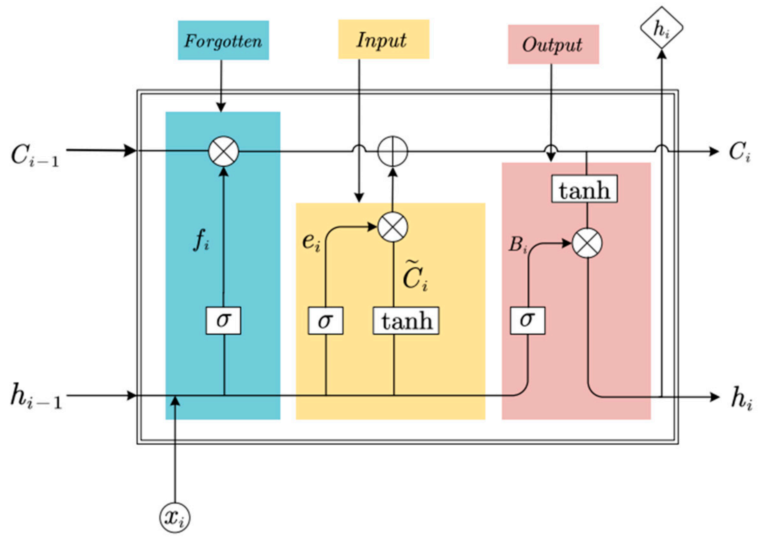

Figure 1. Basic structure diagram of LSTM model.

The basic unit of the LSTM network contains an oblivion gate, an input gate, and an output gate. The oblivion gate determines the oblivion part of the state storage unit by combining the input x_i with the state storage unit C_{i-1} and the intermediate output h_{i-1}, while the input gate transforms x_i by means of the Σ and tanh functions. The associated intermediate output h_i is determined by the updated C_i and the output B_i [32]. The calculation formulas are shown in (22) to (27):

$$f_i = \sigma\left(W_f \cdot [h_{i-1}, x_i] + b_f\right) \tag{22}$$

$$e_i = \sigma(W_e \cdot [h_{i-1}, x_i] + b_e) \tag{23}$$

$$\widetilde{C}_i = \tanh(W_c \cdot [h_{i-1}, x_i] + b_c) \tag{24}$$

$$C_i = f_i * C_{i-1} + e_i * C_i \tag{25}$$

$$B_i = \sigma(W_B \cdot [h_{i-1}, x_i] + b_B) \tag{26}$$

$$h_i = B_i * \tanh(C_i) \tag{27}$$

f_i, e_i, $\widetilde{C}_i$, C_i, and B are the forgetting gate, input gate, new candidate vector, updated cell state, and output gate, respectively, W_f and b_f are the corresponding weight coefficient matrix and bias term, *tanh*, and σ represent the hyperbolic tangent activation function and S-shaped activation function [45]:

$$\tanh(x) = \frac{1 - \exp(-2x)}{1 + \exp(-2x)} \tag{28}$$

$$\sigma(x) = \frac{1}{1 + \exp(-x)} \tag{29}$$

2.2.3. ARIMA-LSTM Model

In order to deal with linear and nonlinear trends in the time series data of chemical composition content before and after weathering of cultural relics, the unique advantages of the ARIMA model in dealing with linear data and the excellent performance of LSTM in dealing with nonlinearity were used [46,47]. First, the artifact chemical content data were processed, and linear prediction results and residual series were obtained with the help of the ARIMA model. Then, the nonlinear factors of the residual series were further analyzed by the LSTM model, and the nonlinear prediction results were obtained. Finally, the linear and nonlinear prediction results were superimposed to obtain the final prediction results for the chemical composition content. According to the decomposition principle of the time series model, it is assumed that the time series $Y = \{y_t, t = 1, 2 \ldots, N\}$ consists of linear and nonlinear components $y_t = x_t + bx_t$. Therefore, the one-dimensional chemical component data are first linearly predicted by the ARIMA model to obtain the linear component x_{tr} and the residual series $\delta_t = y_t + y_{tr}$. Then, the residual series are processed by further nonlinear prediction to obtain the nonlinear component bx_{tr}. Finally, the linear and nonlinear components are combined to obtain the final prediction $y_{tr} = x_{tr} + bx_{tr}$. Root mean square error (RMSE) [48], mean absolute percentage error (MAPE), and R^2 are used to evaluate the performance of the model [49–52].

R^2 is usually taken as [0,1]; the closer R^2 is to 1, the better the fit is, and the equations are as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2} \tag{30}$$

$$\mathrm{MAPE} = \sum_{i=1}^{N}\left|\frac{x_i - y_i}{x_i}\right| \times \frac{100}{N} \tag{31}$$

$$R^2 = \frac{\sum\limits_{i=1}^{N}(y_i - \overline{y})^2 - \sum\limits_{i=1}^{N}(y_i - \hat{y})^2}{\sum\limits_{i=1}^{N}(y_i - \overline{y})^2} \tag{32}$$

In the above equation, x_i is the observed value, y_i is the predicted value, N is the sample size, $\overline{y}$ is the mean of y_i, and $\hat{y}_i$ is the regression fit [32].

2.3. Multiple Linear Regression Model

Multiple linear regression (MLR) is a statistical method that predicts the distribution of the dependent variable by using multiple independent factors [15]. The goal of the MLR model is to establish linear links between independent and dependent characteristics that influence a given event, and it is an extension of classical least squares regression because it employs multiple explanatory factors.

$$y = \alpha_0 + \alpha_1 x_1 + \cdots \alpha_i x_i + \cdots + \alpha_n x_n + \mu_i \tag{33}$$

where y is the dependent variable, $x_1 \cdots x_n$ are the independent variables, α_0 is the y intercept, α_i is the regression coefficient of the ith independent variable, and μ_i is the model error, also known as the residual. The magnitude of the coefficient of determination(R^2)and the squared error(MSE)can be used to assess the predictive performance of the MLR model [15]:

$$MSE = \frac{\Sigma_{j=1}^{n}\left(y_j - \hat{y}_j\right)^2}{n} \tag{34}$$

$$R^2 = 1 - \frac{\sum\limits_{j=1}^{n}\left(y_j - \hat{y}_j\right)^2}{\sum\limits_{j=1}^{n}\left(y_j - \overline{y}_j\right)^2} \tag{35}$$

y_j is the jth parameter after normalization,$\hat{y}_j$ is the jth parameter predicted, $\overline{y}_j$ is the mean of the predicted parameters, and n is the number of samples.

We performed the BP test (Breucsh and Pagan test) and the White test on the perturbation term μ_i to see if there was heteroskedasticity. If the perturbation term is correlated with the independent variable, it may make the regression coefficients of the model inaccurate, thus leading to large errors in the results.

In the BP test, it is assumed that the regression model is $y_i = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_K x_{iK} + \varepsilon_i$, test the following original hypothesis:

$$H_0 : \mathrm{E}\left(\varepsilon_i^2 \middle| x_2, \cdots, x_k\right) = \sigma^2$$

If H_0 is not true, then the conditional variance $\mathrm{E}\left(\varepsilon_i^2 \middle| x_2, \cdots, x_k\right)$ is a function of $(x_2, \cdots, x_k)$ and is called the conditional variance function. The BP test assumes that the conditional variance function is linear:

$$\varepsilon_i^2 = \delta_1 + \delta_2 x_{i2} + \cdots + \delta_K x_{iK} + u_i \tag{36}$$

The original hypothesis can be simplified to:

$$H_0 : \delta_2 = \cdots = \delta_K = 0$$

If we assume that H_0 is true, we can show that ε_i has no correlation with the independent variable x_{iK}; that is, there is no autocorrelation, and the perturbation term has no heteroscedasticity. Since the perturbation term ε_i is not observable, the residual squared e_i^2 is used for auxiliary regression of the explanatory variable:

$$e_i^2 = \delta_1 + \delta_2 x_{i2} + \cdots + \delta_K x_{iK} + error_i \tag{37}$$

nR^2 statistics were used:

$$nR^2 \xrightarrow{d} x^2(K-1)$$

R^2 is the R^2 of auxiliary regression. The difference between the White test and the BP test lies in that when the White test carries out auxiliary regression, there are x_{iK} square terms and cross terms in Equation (37), so the BP test can be regarded as a special case of the White test.

In addition, we tested the model for multicollinearity, and the variance inflation factor VIF was used to eliminate the influence factors with multicollinearity, which improved the accuracy of the model:

Assuming that there are k independent variables, then the variance inflation factor $VIF_n = \frac{1}{1-R^2_{1-k/n}}$, $R^2_{1-k/n}$ is the goodness of fit obtained by regressing the n-th independent variable as the dependent variable on the remaining $k-1$ independent variables; the larger the VIF_n, the greater the correlation between the n-th variable and the other variables [53]. If VIF_n is greater than 10, there is strict multicollinearity between the variables.

3. Material and Experiment

3.1. Data Pre-Processing

This study is based on the data related to ancient glassware provided by the official website of the 2022 China Student Mathematical Modeling Competition [14]. The glass sampling points are discussed separately by two systems: lead-barium glass and high-potassium glass. The data gives the proportion of the chemical composition of the sampling points of this batch of artifacts, which is characterized by composition, that is, the data of the proportion of the content of each chemical component of the cumulative sum should be found 100%, but may be due to detection means or contain various types of impurities and other reasons, resulting in the proportion of its corresponding components of the cumulative sum of the non-100% situation. Thus, in this study, the data with the sum of components between 85% and 105% were stored as valid data, and the data with severe weathering of the glass were excluded to eliminate the influence of outliers on the model results. The results are shown in Table A1.

3.2. Experimental Procedure

This paper focuses on three improved joint model algorithms, Daen-LR, ARIMA-LSTM and MLR. The experimental software environment Matlab 2021b, SPSS, Stata were used to analyze the identification of ancient glasses.

First, in this paper, the obtained pre-processed data set is used to find the relationship between the chemical composition content and weathering at its sampling points after glass classification by building an improved logistic regression model based on a double adaptive elastic net. Then, by using the ARIMA-LSTM model, we predict the content of chemical components contained in the two glasses before weathering and obtain the correlation curves of chemical components before and after weathering. Finally, based on the results obtained above, this paper uses a multiple linear regression model to predict the type of unknown glass and judges the accuracy and efficiency of the model by testing whether there is heteroskedasticity in the perturbation terms, multicollinearity, and the degree of fit between the experimental and actual values of the model. The flow chart is shown in Figure 2.

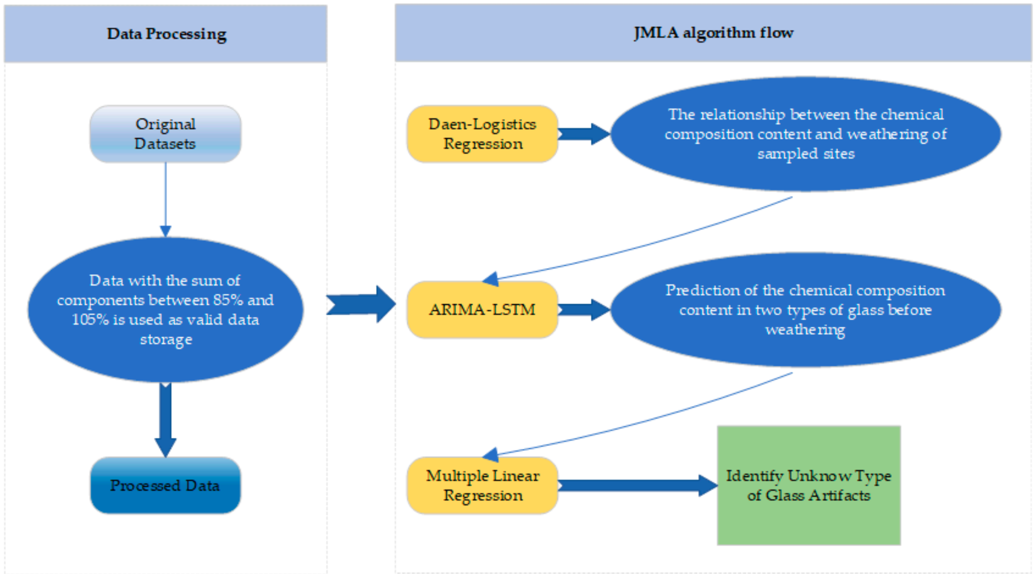

Figure 2. Processing flow chart.

4. Process and Result

4.1. Relationship between Glass Weathering and Its Chemical Composition Based on an Improved Logistic Regression Model with Double Adaptive Elastic Net

Based on the data in Table A1, we use Matlab and SPSS to conduct modeling and calculation of the Daen-Logistic Regression model. The dependent variable here is a dichotomous variable (i.e., weathered and unweathered states), and the content of various chemical components is set as the independent variable. The double adaptive elastic net model can determine the classification results of weathered or unweathered glass under different conditions for each independent variable, which can avoid the variability of the results when the independent variables are selected differently, make the classification more adaptive, and reduce the influence of strongly correlated variables on other variables. The calculation gives the following in Table 1:

Table 1. Return coefficients of chemical components in high potassium and lead-barium glasses β and significance *p*-value.

Regression Coefficient of High Potassium Glass Types β		Significance *p*-Value of High Potassium Glass type (P > |t|)		Regression Coefficient of Lead-Barium Glass Types β		Significance *p*-Value of Lead-Barium Glass Type (P > |t|)	
β_0	15.301	P_0	0.000	β_0	31.077	P_0	0.000
β_1	2.718	P_1	0.099	β_1	7.788	P_1	0.005
β_2	12.318	P_2	0.000	β_2	1.016	P_2	0.313
β_3	5.410	P_3	0.020	β_3	8.320	P_3	0.004
β_4	7.199	P_4	0.007	β_4	0.039	P_4	0.843
β_5	7.051	P_5	0.008	β_5	2.678	P_5	0.102
β_6	4.751	P_6	0.029	β_6	0.222	P_6	0.637
β_7	1.072	P_7	0.300	β_7	1.293	P_7	0.255
β_8	2.629	P_8	0.101	β_8	25.165	P_8	0.000
β_9	1.792	P_9	0.181	β_9	0.831	P_9	0.362
β_{10}	2.629	P_{10}	0.105	β_{10}	13.764	P_{10}	0.000
β_{11}	3.142	P_{11}	0.076	β_{11}	3.702	P_{11}	0.054
β_{12}	0.451	P_{12}	0.502	β_{12}	0.161	P_{12}	0.688
β_{13}	1.490	P_{13}	0.222	β_{13}	0.942	P_{13}	0.332

From the table of high potassium glass type, it can be seen that the values of two chemical components, SiO_2 and K_2O, are relatively large, and the values of the significance *p*-value are less than $p = 0.1$, so these two chemical components have the greatest influence on whether the surface of high potassium glass is weathered or not. From the table of lead-barium glass type, it can be seen that the values of three chemical components, SiO_2, PbO, and P_2O_5, are relatively large, and the values of the significance *p*-value are less than $p = 0.1$, so these two chemical components have the greatest influence on whether the surface of high potassium glass is weathered or not.

4.2. Prediction of the Chemical Content of Glass before Weathering Based on the ARIMA-LSTM Model

By solving 4.1, we obtained the results of the relationship between glass weathering and its chemical composition, and using this relationship, we screened 14 chemical elements in two respective types, high potassium and lead-barium, respectively. For high potassium glasses, we have chosen to retain both SiO_2 and K_2O chemical components. For lead-barium glasses, we chose to retain three chemical components: SiO_2, PbO, and P_2O_5; all of them have relatively complete data and have strong correlations for modeling analysis.

In order to make the model better identify the patterns in the data, the outliers with large deviations are first eliminated. SPSS 24 software was used to detect three abnormal data values with an additive or transient state, and the existence of such outliers would lead to accidental results in the model, leading to wrong conclusions. Taking the SiO_2 content of high potassium glass and lead-barium glass as examples, the outliers of both are shown in Table 2.

Table 2. Outlier of SiO_2 component content.

Relic Number	Outlier Type		Estimates	S.E.	t	Significance
High potassium 02	Additive		23.850	3.478	6.858	0.000
High potassium 13	Transient	Magnitude	16.260	3.478	4.675	0.001
		Decay factor	0.829	0.266	3.114	0.011
Lead barium 05	Transient	Magnitude	31.926	4.569	6.988	0.000
		Decay factor	0.964	0.013	76.495	0.000

Through the analysis and calculation in Matlab and SPSS, we tested and fitted the data values of all the chemical composition contents changing with the time series and established the ARIMA-LSTM prediction curve model. We found that the parameters of ARIMA(2,1,0)-LSTM can obtain the maximum likelihood value of the model, and the normalized BIC [54] values of 3.160 and 4.160 for the SiO_2 component content in high potassium glass and lead-barium glass, for example, are the smallest values among the parameters. In addition, the smooth R^2 values of the model are 0.960 and 0.934, both close to 1, and both p-values are 0.000, both less than 0.05, so it can be considered that the results of the model are significantly reasonable and can fit well with the prediction model (Table 3).

Table 3. Parameters of the ARIMA-LSTM model for SiO_2 component content in two types of glass.

Type \ Fitting Statistics	Stationary R^2	R^2	RMSE	MAPE	MaxAPE	MAE	MaxAE	Normalized BIC
High potassium glass	0.960	0.960	1.330	1.411	8.842	2.177	6.130	3.160
Lead-barium glass	0.934	0.934	1.523	2.105	10.624	4.001	11.250	4.160

After the initial completion of the estimated time series model based on the chemical composition content, a white noise test of the residuals is required. If the residuals are white noise, then it can indicate that the selected model can identify the laws of the time series data, that is, the model is acceptable; if the residuals are not white noise, then it means that there is still some information not identified; at this time, the model parameters need to be revised to continue to identify this part of the information. The study used Ljung and Box's Q test to determine whether the residuals are white noise [55,56]:

Assuming that the residual $\{\epsilon_t\}$ is a white noise sequence, then $\rho_s = \begin{cases} 1, s = 0 \\ 0, s \neq 0 \end{cases}$, the autocorrelation coefficient of the sample, is:

$$r = \hat{\rho_s} = \frac{\sum_{t=s+1}^{T} (x_t - \overline{x})(x_{t-s} - \overline{x})}{\sum_{t=1}^{T} (x_t - \overline{x})^2} \tag{38}$$

In $H_0 : \rho_1 = \rho_2 = \cdots = \rho_S = 0$, $H_1 : \rho_i (i = 1, 2, \ldots, s)$ at least one is not 0. In the case that H_0 holds, the statistic $Q = T(T+2) \sum_{k+1}^{s} \frac{r_k^2}{T-k} \sim X_{s-n}^2$, from which the *p*-value can be calculated, and if the *p*-value is less than 0.05, then the original hypothesis is rejected, indicating that the model is not fully identified and the model parameters need to be modified.

Through the model statistics, the *p*-values of the Ljung and Box's Q test for SiO_2 content of high potassium glass and lead-barium glass are 0.889 and 0.744, respectively, both of which are greater than 0.05, i.e., we cannot reject the original hypothesis, and we can assume that the residuals are white noise sequences and the model can be fully identified. Figure 3 shows that the autocorrelation coefficients and partial autocorrelation coefficients of all lag orders are not significantly different from 0 [57,58].

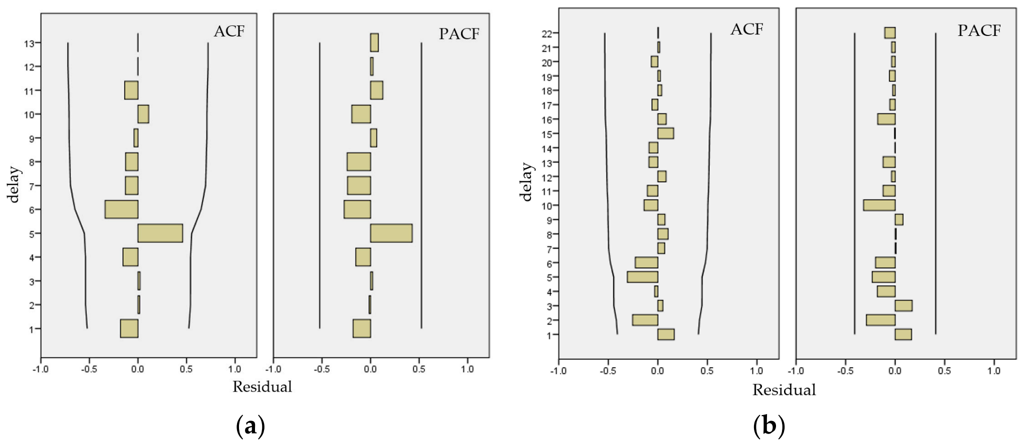

Figure 3. (**a**) ACF and PACF patterns of SiO_2 content in high potassium; (**b**) ACF and PACF patterns of SiO_2 content in lead-barium glasses.

By the same method, the fitting coefficients of all mathematical models of the measured chemical composition contents were obtained. In the category of high potassium glass, the R^2 values of SiO_2 and K_2O were 0.960 and 0.969, respectively. In the category of lead-barium glass, the R^2 values of SiO_2, P_2O_5, and PbO are 0.934, 0.951, and 0.948, respectively. Finally, the corresponding prediction model curve is drawn, from which the correlation of chemical composition content before and after weathering can be clearly seen, as shown in Figure 4. The blue curve represents the actual value of chemical content changing with time after weathering, while the yellow curve represents the fitting value. The fitting degree of both represents the superiority of the model's performance. The red curve represents the predicted value of component content over time before weathering. It can be seen that the ARIMA (2,1,0)-LSTM model shows the correlation of chemical composition contents before and after weathering, reduces the interference of "weathering" factors on glass classification, and improves the accuracy of subsequent glass classification.

4.3. Identifying Unknown Artifact Types Based on Multiple Linear Regression Model

Through the results of Sections 4.1 and 4.2, we conducted chemical content testing and analysis on a batch of newly excavated glass artifacts, as shown in Table 4, and judged the categories to which they belonged by the correlation of related elements and weathering effects. Firstly, the chemical elements with significant correlation in each category were initially screened out by the statistical law of chemical element content, and the multiple linear regression equation between elements and categories was established to find out the experimental values of categories and make errors and fits with the actual values of categories, and to verify the accuracy of the model.

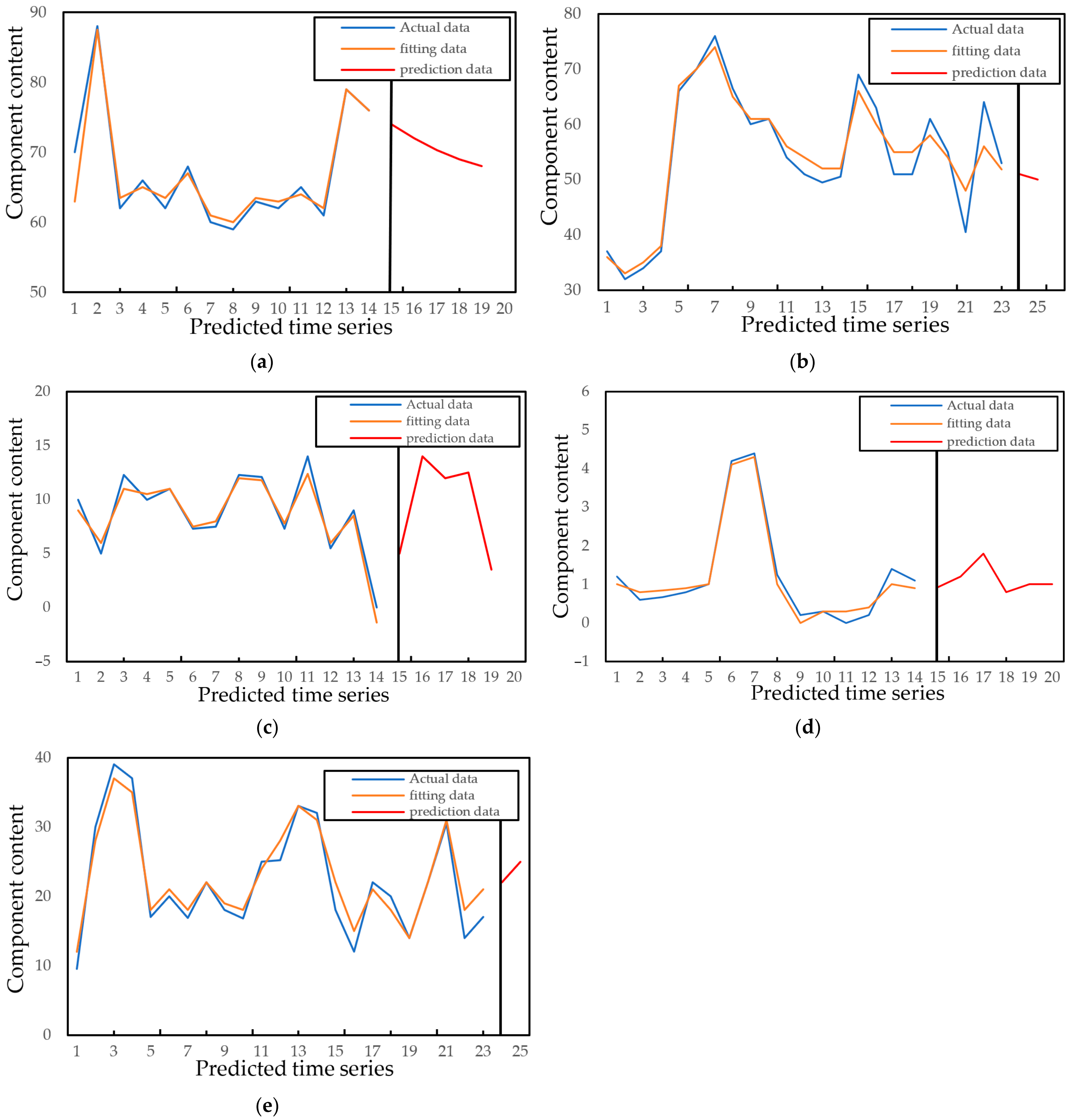

Figure 4. Correlation curve of chemical element content before and after weathering: (**a**) (High potassium) SiO_2; (**b**) (lead barium) SiO_2; (**c**) (High potassium) K_2O; (**d**) (lead barium) P_2O_5; (**e**) (lead barium) PbO.

According to the classification rules of chemical content and surface weathering, it can be initially concluded that K_2O, CaO, MgO, Al_2O_3, FeO, PbO, BaO, and P_2O_5 have strong correlations with surface weathering and categories, while the remaining elements have weak correlations, so the remaining elements can be deleted. In addition, because the chemical element contents of the three heavily weathered glass artifacts are very different from other contents, which will have a large impact on the analysis of the model, they are treated as outliers. In the multiple linear regression equation, the qualitative data should be set as dummy variables, so the qualitative variables (unweathered and weathered) in the surface weathering independent variable *Suw* can be set as quantitative variables

(0 and 1), and the qualitative variables (high-potassium and lead-barium) in the discriminatory category dependent variable y_i can be set as quantitative variables (A and B), and the following multiple linear regression equation can be established as:

$$\begin{aligned} y_i = \quad & \alpha_0 + \alpha_1 x_{Sii} + \alpha_2 x_{Ki} + \alpha_3 x_{Cai} + \alpha_4 x_{Mgi} + \alpha_5 x_{Ali} \\ & + \alpha_6 x_{Fei} + \alpha_7 x_{Pbi} + \alpha_8 x_{Bai} + \alpha_9 x_{Pi} + \beta Suw_i + \mu_i \end{aligned} \tag{39}$$

$$\begin{gathered} Suw_i = 1 \text{ denotes the } i - \text{th weathering sample} \\ Suw_i = 0 \text{ denotes the } i - \text{th unweathered sample} \\ E(y \mid Suw = 1 \; and \; other \; independent \; variables) = \beta \times 1 + m(Constants) \\ E(y \mid Suw = 0 \; and \; other \; independent \; variables) = \beta \times 0 + m(Constants) \end{gathered}$$

Table 4. Chemical Composition of Unclassified Cultural Relics.

Relic Number	**A1**	**A2**	**A3**	**A4**	**A5**	**A6**	**A7**	**A8**
Surface Weathering	**No**	**Yes**	**No**	**No**	**Yes**	**Yes**	**Yes**	**No**
SiO_2	78.45	37.75	31.95	35.47	64.29	93.17	90.83	51.12
Na_2O	0.00	0.00	0.00	0.00	1.2	0.00	0.00	0.00
K_2O	0.00	0.00	1.36	0.79	0.37	1.35	0.98	0.23
CaO	6.08	7.63	7.19	2.89	1.64	0.64	1.12	0.89
MgO	1.86	0.00	0.81	1.05	2.34	0.21	0.00	0.00
Al_2O_3	7.23	2.33	2.93	7.07	12.75	1.52	5.06	2.12
Fe_2O_3	2.15	0.00	7.06	6.45	0.81	0.27	0.24	0.00
CuO	2.11	0.00	0.21	0.96	0.94	1.73	1.17	9.01
PbO	0.00	34.3	39.58	24.28	12.23	0.00	0.00	21.24
BaO	0.00	0.00	4.69	8.31	2.16	0.00	0.00	11.34
P_2O_5	1.06	14.27	2.68	8.45	0.19	0.21	0.13	1.46
SrO	0.03	0.00	0.52	0.28	0.21	0.00	0.00	0.31
SnO_2	0.00	0.00	0.00	0.00	0.49	0.00	0.00	0.00
SO_2	0.51	0.00	0.00	0.00	0.00	0.00	0.11	2.26

The joint significance test indicators for the F-statistic [59,60] for the above model results are as follows:

F(10,55) represents the F joint statistic test value of 51.41, the confidence interval is 95%, and the original hypothesis H_0 is: $a1 = a2 = a3 = \cdots = a9 = \beta = 0$. From Table 5, we can see that the p-value is 0, p is less than 0.05, and at this time the original hypothesis is rejected. We have reason to believe that the correlation coefficient is significantly different from 0, so we can consider this model to be useful. The regression coefficients and corresponding p-values for the variables of interest can be derived as in Table 6. Only when the p-value is less than 0.05, we consider it significant, and the regression coefficient is credible at this point, so we can use the regression coefficients corresponding to K_2O, Al_2O_3, PbO, BaO, and *Suw* (the dummy variable "weathering"), and the larger the absolute value of the regression coefficients, the greater the effect on the dependent variable.

Table 5. Indicators for joint significance testing of F-statistics.

F(10,55)	51.41
Prob > F	0.0000
R-squared	0.9633
Adj R-squared	0.9558

Table 6. Regression coefficient β and significance *p*-value.

Type	SiO_2	K_2O	CaO	MgO	Al_2O_3	Fe_2O_3
Coef.	−0.014	0.048	−0.001	0.004	−0.299	0.035
P > ∣t∣	0.749	0.000	0.965	0.919	0.001	0.105
Type	**PbO**	**BaO**	**P_2O_5**	**SUW**	**y_i**	**_cons**
Coef.	−0.166	−0.161	−0.018	0.352	0.000	0.746
P > ∣t∣	0.001	0.021	0.078	0.000	0.000	0.082

We can derive the multiple linear regression equation for glass artifact class, chemical element content, and weathering type as follows:

$$\hat{y}_i = 0.7458 + 0.0480x_k - 0.0299x_{Al} - 0.0165x_{Pb} - 0.0161x_{Ba} + 0.3517\,\mathrm{Suw} \tag{40}$$

4.3.1. Testing for the Presence of Heteroskedasticity in the Perturbation Term

Perturbation term μ_i is unobservable and requires certain conditions to be met. Our model defaults to a spherical perturbation term, which generally has to satisfy "no autocorrelation" and "homoskedasticity" because if the perturbation term is "correlated with the independent variable", i.e., endogenous, it will make the correlation regression coefficient inaccurate; if there is "heteroskedasticity", it will cause the hypothesis test statistic we constructed to be invalid, and the *OLS* estimator cannot be treated as the optimal linear unbiased estimator [61]. Therefore, we performed the BP test and White test on the perturbation term to verify the presence of heteroskedasticity, as shown in Table 7 [62–64].

Table 7. Results of the BP test and White test.

BP test	Prob > chi2	0.1725
White test	Prob > chi2	0.4095

The above two hypotheses were tested for heteroskedasticity, and the original hypothesis *H*0 was that there is no heteroskedasticity in the perturbation term. However, the *p*-value is greater than 0.05, so *H*0 is accepted, and we can assume that there is no heteroskedasticity in the perturbation term.

4.3.2. Testing for Multicollinearity

If the data matrix X does not satisfy the column rank, i.e., a variable can be linearly expressed by other explanatory variables, then there is "strict multicollinearity", Stata software was used to calculate the VIF of each variable, and the test results were as follows Table 8:

Table 8. Results of the variance inflation factor analysis.

Variable	VIF
SiO_2	27.28
PbO	21.06
BaO	6.36
K_2O	5.11
P_2O_5	2.75
CaO	2.68
MgO	1.83
Al_2O_3	1.82
Fe_2O_3	1.63
Mean Value	7.84

It is generally believed that when VIF > 10, the regression equation has severe multicollinearity; SiO_2 and PbO both exceed 10, but the *p*-value of SiO_2 is higher than 0.05, which is not significant, so its coefficient is not considered in the equation model. PbO, although VIF exceeds 10, the p-value is lower than 0.05 because the coefficient is still significant with variance inflation; if there is no multicollinearity, the regression coefficients would be more significant.

4.3.3. Testing the Fit of the Experimental and Actual Values of the Model and Identifying the Unknown Artifact Types

Due to the fact that the dependent variable is the category of glass artifacts, there are only two categories: high-potassium and lead-barium, so it can be treated as a 1-0 variable. If the experimental value is close to 1, then it is considered the high potassium category; if the experimental value is close to 0, then it is considered the lead-barium category. From Figure 5 and Table 9, it can be seen that the 66 samples fit very well, almost no chance data occur, and the identification results of 8 unknown cultural relic types are completely consistent with reality, so it can be considered that the predicted value of this multiple linear regression equation is quite accurate.

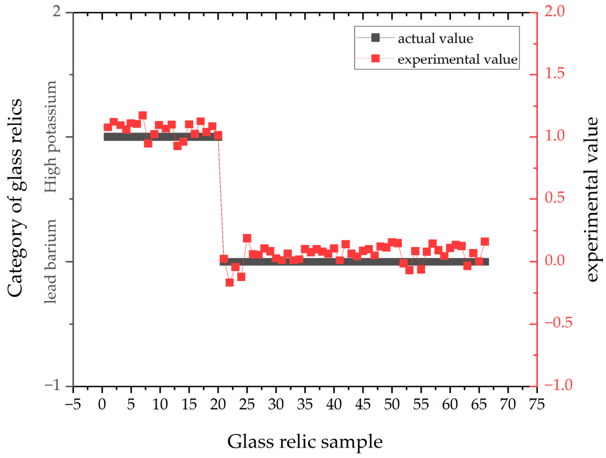

Figure 5. Fitting curve of multiple linear regression equation.

Table 9. Identification results.

Relic Number	A1	A2	A3	A4	A5	A6	A7	A8
Identification type	High potassium	Lead barium	Lead barium	Lead barium	Lead barium	High potassium	High potassium	Lead barium

4.4. Comparision of Different Models

To demonstrate the superiority of the proposed method, we compare the proposed joint algorithm with similar decision and classification algorithms like Decision Trees (DT), Random Forests (RF) [65], Support Vector Machines (SVM), Random Forests based on calssification and regression tree (CART-RF) [66–69]. We did not use any pre-trained models, but trained each model from scratch. When we select the parameters of traditional machine learning algorithm, we take into account the number of data features and avoid overfitting, as we can see in Table 10. Then we perform experimental simulations of these models to be compared as well as the model proposed in this paper using Matlab. The results are presented in the following Table 11. In the classification results, this study uses common evaluation indicators to judge the superiority of the model: Train Acc, Test Acc, Precision, Recall, and F_1 Score. TP, TN, FP, FN are required to explain the above indicators, so confusion matrix is introduced, as shown in Figure 6. The specific performance is described as follows:

Table 10. Selected parameters of each algorithm.

Algorithms	Parameters
DT	Node splitting evaluation criteria = gini Feature division point selection criteria = random Minimum samples for internal node splitting = 2 Minimum samples in leaf nodes = 1 Maximum leaf nodes = 2 Maximum depth of the tree = 15

Table 10. *Cont.*

Algorithms	Parameters
RF	Node split evaluation criterion = gini Number of decision trees = 5 Minimum samples in leaf nodes = 1 Maximum depth of the tree = 15 Maximum leaf nodes = 2
CART-RF	Node split evaluation criterion = gini Number of decision trees = 6 Minimum samples in leaf nodes = 3 Maximum depth of the tree = 15 Maximum leaf nodes = 2
SVM	kernel = 'rbf' C = 20 $\gamma = 2.00$
JMLA	Qualitative variable—weathered: 1 Qualitative variable—unweathered: 0 The number of autoregressive terms: 2 The number of sliding average terms: 0 The number of differences needed to make it a smooth series: 1

Table 11. Results of model experiments.

Algorithms	Train Acc	Test Acc	Precision	Recall	F_1 Score
DT	0.862	0.873	0.806	0.791	0.798
RF	0.958	0.924	0.872	0.866	0.869
CART-RF	0.962	0.951	0.929	0.941	0.935
SVM	0.850	0.909	0.869	0.830	0.849
JMLA	0.979	0.976	0.975	0.976	0.975

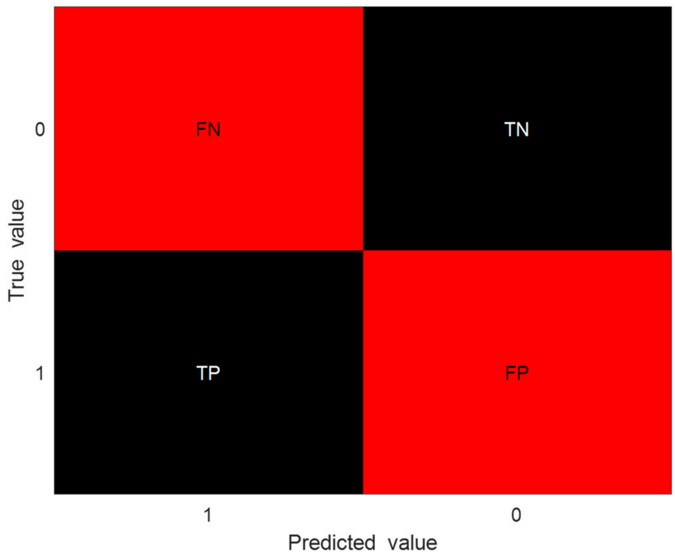

Figure 6. Confusion matrix (1: high potassium, 0: lead barium).

TP (True Positive): The true value of the data is high potassium, and the predicted value is also high potassium.
TN (True Negative): The true value of the data is lead barium, and the predicted value is also lead barium.
FP (False Positive): The true value of the data is high potassium, but it is incorrectly predicted as lead barium.

FN (False Negative): The true value of the data is lead barium, but it is incorrectly predicted as high potassium.

Accuracy is the simplest and most clear index for evaluating classification models, but it is a good measurement standard only when the proportion of samples in each category of the data set is fairly balanced, as shown in Equation (41):

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{41}$$

Precision represents the proportion of samples that are actually positive in the predicted positive example. As shown in Equation (42):

$$\text{Precision} = \frac{TP}{TP + FP} \tag{42}$$

Recall represents the proportion of the actual number of positive samples in the total positive samples among the predicted positive samples. As shown in Equation (43):

$$\text{Recall} = \frac{TP}{TP + FN} \tag{43}$$

F_1 Score is a weighted average of accuracy rate and recall rate, which is a synthesis of both. The value of the F_1 Score determines the robustness of the model. It can be considered that the higher F_1 is, the more stable the model is. As shown in Equation (44):

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{44}$$

5. Discussion

From the results of the comparison experiments, it is not difficult to see that the joint algorithm proposed in this paper shows notable advantages in all performance parameters. As shown in Table 11, in the indicators of Train Acc, Test Acc, Precision, Recall, and F_1 Score, we can find that the difference values of JMLA's performance over the past optimal algorithm model are +0.017, +0.025, +0.046, +0.035, and +0.040, respectively. We improve common machine learning algorithms and combine them with deep learning models to make the classification results more accurate, which provides a new idea for the study of the classification of ancient cultural relics.

Considering the accuracy of the JMLA algorithm in classification results and excellent evaluation indexes, this study believes that the model proposed in this paper is suitable for providing more in-depth research ideas for the classification of ancient cultural relics. The algorithm ideas in this paper can also be applied to other related fields, such as the data analysis of nutrient elements in food, the influence of air oxidation degree on nutrient elements, the classification of water pollution degree, etc. However, there is still room for improvement in the joint algorithm to address its high computational complexity and formula complexity. Compared with the existing algorithm, the calculation cost of JMLA is higher, and the formula is more complex. Further reducing algorithm complexity and better unifying the above three algorithms will be the focus of future research.

6. Conclusions

In this paper, we propose a joint Daen-LR, ARIMA-LSTM, and MLR machine learning algorithm (JMLA). Firstly, we combine a double adaptive elastic network with a traditional logistic model to select variables that have both Oracle and adaptive classification characteristics. These two characteristics eliminate the influence of different categories on the inconsistent selection of important independent variables and the influence of strong-correlation independent variables on the interference of weak independent variables. Secondly, we combine the deep learning model (LSTM) with the ARIMA time series model so that it can handle both linear and nonlinear trends. By calculating the ARIMA-LSTM model, we establish the correlation curve of chemical composition before and after weathering and predict the change in chemical composition with weathering. Thirdly, we combine the data processed by the above two improved methods with the multiple linear regression model to classify the unknown glass relics.

The experimental results show that the accuracy of the JMLA model on the train set is 97.9%, and the accuracy of the JMLA model on the test set is 97.6%. In addition, we compared JMLA with similar classifiers, and the results were shown in Train Acc, Test Acc, Precision, Recall, and F_1

Score indexes. The difference values of JMLA's performance over the past optimal algorithm model are +0.017, +0.025, +0.046, +0.035, and +0.040, respectively. These data show that the JMLA model has better performance than other classification models without changing the structure of similar classification models and under the same experimental conditions. The classification accuracy of the JMLA model is higher than other models, especially for large glass relics with more chemical elements and a harsh environment.

This processing method is practical and reliable in the direction related to the composition analysis and identification of cultural heritage. The application of this method is expected to improve the accuracy of the classification of cultural relics by archaeologists and can effectively reduce the impact of identification difficulties caused by factors such as harsh burial environments. It helps us to have a deeper understanding of the exchange, penetration, and development of ancient Eastern and Western cultures.

In addition, the future research directions of this study can be summarized as follows:

1. Algorithm optimization. The processing method uses a variety of machine learning algorithms that effectively combine the advantages of each algorithm with high practicality and feasibility and a good fitting effect. However, this model is only combined with an LSTM deep learning neural network, which can be combined with more advanced deep learning models in the future so as to improve the accuracy and efficiency of classification.
2. Reduce model calculation costs and formula complexity. Although the classification accuracy of the JMLA model is very high, the calculation time is relatively long compared with other models, and the formula is relatively complex, which is also a pain point for the JMLA model. Therefore, reducing the calculation amount and better integrating the three models will be the focus of future research.
3. Application prospects of this data processing method. In spite of its application in the direction of heritage composition analysis and identification, it is expected to be applied in the areas of health, food safety, and environmental protection, for example: analysis and classification of chemical constituents of tobacco; composition analysis of nutritional composition in food; classification and monitoring of pollutant composition in air, etc.

Author Contributions: Conceptualization, Z.-X.L. and P.-S.L.; data curation, P.-S.L.; formal analysis, Z.-X.L.; investigation, J.-H.L.; methodology, Z.-X.L.; project administration, Z.-H.Y., Y.-P.M. and H.-H.W.; resources, G.-Y.W.; software, P.-S.L.; supervision, G.-Y.W. and H.-H.W.; validation, J.-H.L., Z.-H.Y. and Y.-P.M.; visualization, J.-H.L.; writing—original draft, Z.-X.L. and P.-S.L.; writing—review and editing, G.-Y.W., J.-H.L. and Z.-H.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (62203332/22008050) and the Natural Science Foundation of Hebei Province (B2022202008).

Institutional Review Board Statement: This study did not require ethical approval.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are provided in the 2022 China Undergraduate Mathematical Contest in Modeling. The authors do not have permission to share the data.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. The result graph after glass pre-processing.

Glass Sampling Points	SiO_2	Na_2O	K_2O	CaO	MgO	Al_2O_3	Fe_2O_3	CuO	PbO	BaO	P_2O_5	SrO	SnO_2	SO_2
HPNP01	69.33	0	9.99	6.32	0.87	3.93	1.74	3.87	0	0	1.17	0	0	0.39
HPNP03(1)	87.05	0	5.19	2.01	0	4.06	0	0.78	0.25	0	0.66	0	0	0
HPNP03(2)	61.71	0	12.37	5.87	1.11	5.5	2.16	5.09	1.41	2.86	0.7	0.1	0	0
HPNP04	65.88	0	9.67	7.12	1.56	6.44	2.06	2.18	0	0	0.79	0	0	0.36
HPNP05	61.58	0	10.95	7.35	1.77	7.5	2.62	3.27	0	0	0.94	0.06	0	0.47

Table A1. *Cont.*

Glass Sampling Points	SiO_2	Na_2O	K_2O	CaO	MgO	Al_2O_3	Fe_2O_3	CuO	PbO	BaO	P_2O_5	SrO	SnO_2	SO_2
HPNP06(1)	67.65	0	7.37	0	1.98	11.15	2.39	2.51	0.2	1.38	4.18	0.11	0	0
HPNP06(2)	59.81	0	7.68	5.41	1.73	10.05	6.04	2.18	0.35	0.97	4.5	0.12	0	0
HPWP07	92.63	0	0	1.07	0	1.98	0.17	3.24	0	0	0.61	0	0	0
HPWP09	95.02	0	0.59	0.62	0	1.32	0.32	1.55	0	0	0.35	0	0	0
HPWP10	96.77	0	0.92	0.21	0	0.81	0.26	0.84	0	0	0	0	0	0
HPWP12	94.29	0	1.01	0.72	0	1.46	0.29	1.65	0	0	0.15	0	0	0
HPNP13	59.01	2.86	12.53	8.7	0	6.16	2.88	4.73	0	0	1.27	0	0	0
HPNP14	62.47	3.38	12.28	8.23	0.66	9.23	0.5	0.47	1.62	0	0.16	0	0	0
HPNP15	61.87	3.21	7.44	0	1.02	3.15	1.04	1.29	0.19	0	0.26	0	0	0
HPNP16	65.18	2.1	14.52	8.27	0.52	6.18	0.42	1.07	0.11	0	0	0.04	0	0
HPNP17	60.71	2.12	5.71	0	0.85	0	1.04	1.09	0.19	0	0.18	0	0	0
HPNP18	79.46	0	9.42	0	1.53	3.05	0	0	0	0	1.36	0.07	2.36	0
HPNP21	76.68	0	0	4.71	1.22	6.19	2.37	3.28	1	1.97	1.1	0	0	0
HPWP27	92.72	0	0	0.94	0.54	2.51	0.2	1.54	0	0	0.36	0	0	0
HPWP22	92.35	0	0.74	1.66	0.64	3.5	0.35	0.55	0	0	0.21	0	0	0
LBNP20	37.36	0	0.71	0	0	5.45	1.51	4.78	9.3	23.55	5.75	0	0	0
LBNP23	53.79	7.92	0	0.5	0.71	1.42	0	2.99	16.98	11.86	0	0.33	0	0
LBNP24	31.94	0	0	0.47	0	1.59	0	8.46	29.14	26.23	0.14	0.91	0	0
LBNP25	50.61	2.31	0	0.63	0	1.9	1.55	1.12	31.9	6.65	0.19	0.2	0	0
LBWP26	19.79	0	0	1.44	0	0.7	0	10.57	29.53	32.25	3.13	0.45	0	1.96
LBWP08	20.14	0	0	1.48	0	1.34	0	10.41	28.68	31.23	3.59	0.37	0	2.58
LBWP19	29.64	0	0	2.93	0.59	3.57	1.33	3.51	42.82	5.35	8.83	0.19	0	0
LBWP11	33.59	0	0.21	3.51	0.71	2.69	0	4.93	25.39	14.61	9.38	0.37	0	0
LBWP02	36.28	0	1.05	2.34	1.18	5.73	1.86	0.26	47.43	0	3.57	0.19	0	0
LBNP28	68.08	0	0.26	1.34	1	4.7	0.41	0.33	17.14	4.04	1.04	0.12	0.23	0
LBNP29	63.3	0.92	0.3	2.98	1.49	14.34	0.81	0.74	12.31	2.03	0.41	0.25	0	0
LBNP30(1)	34.34	0	1.41	4.49	0.98	4.35	2.12	0	39.22	10.29	0	0.35	0.4	0
LBNP30(2)	36.93	0	0	4.24	0.51	3.86	2.74	0	37.74	10.35	1.41	0.48	0.44	0
LBNP31	65.91	0	0	1.6	0.89	3.11	4.59	0.44	16.55	3.42	1.62	0.3	0	0
LBNP32	69.71	0	0.21	0.46	0	2.36	1	0.11	19.76	4.88	0.17	0	0	0
LBNP33	75.51	0	0.15	0.64	1	2.35	0	0.47	16.16	3.55	0.13	0	0	0
LBWP34	35.78	0	0.25	0.78	0	1.62	0.47	1.51	46.55	10	0.34	0.22	0	0
LBNP35	65.91	0	0	0.38	0	1.44	0.17	0.16	22.05	5.68	0.42	0	0	0
LBWP36	39.57	2.22	0.14	0.37	0	1.6	0.32	0.68	41.61	10.83	0.07	0.22	0	0
LBNP37	60.12	0	0.23	0.89	0	2.72	0	3.01	17.24	10.34	1.46	0.31	0	3.66
LBWP38	32.93	1.38	0	0.68	0	2.57	0.29	0.73	49.31	9.79	0.48	0.41	0	0
LBWP39	26.25	0	0	1.11	0	0.5	0	0.88	61.03	7.22	1.16	0.61	0	0
LBWP40	16.71	0	0	1.87	0	0.45	0.19	0	70.21	6.69	1.77	0.68	0	0
LBWP41	18.46	0	0.44	4.96	2.73	3.33	1.79	0.19	44.12	9.76	7.46	0.47	0	0
LBNP42(1)	51.26	5.74	0.15	0.79	1.09	3.53	0	2.67	21.88	10.47	0.08	0.35	0	0
LBNP42(2)	51.33	5.68	0.35	0	1.16	5.66	0	2.72	20.12	10.88	0	0	0	0
LBWP43(1)	12.41	0	0	5.24	0.89	2.25	0.76	5.35	59.85	7.29	0	0.64	0	0
LBWP43(2)	21.7	0	0	6.4	0.95	3.41	1.39	1.51	44.75	3.26	12.83	0.47	0	0

Table A1. *Cont.*

Glass Sampling Points	SiO_2	Na_2O	K_2O	CaO	MgO	Al_2O_3	Fe_2O_3	CuO	PbO	BaO	P_2O_5	SrO	SnO_2	SO_2
LBNP44	60.74	3.06	0.2	2.14	0	12.69	0.77	0.43	13.61	5.22	0	0.26	0	0
LBNP45	61.28	2.66	0.11	0.84	0.74	5	0	0.53	15.99	10.96	0	0.23	0	0
LBNP46	55.21	0	0.25	0	1.67	4.79	0	0.77	25.25	10.06	0.2	0.43	0	0
LBNP47	51.54	4.66	0.29	0.87	0.61	3.06	0	0.65	25.4	9.23	0.1	0.85	0	0
LBWP48	53.33	0.8	0.32	2.82	1.54	13.65	1.03	0	15.71	7.31	1.1	0.25	1.31	0
LBNP49	28.79	0	0	4.58	1.47	5.38	2.74	0.7	34.18	6.1	11.1	0.46	0	0
LBWP49	54.61	0	0.3	2.08	1.2	6.5	1.27	0.45	23.02	4.19	4.32	0.3	0	0
LBWP50	17.98	0	0	3.19	0.47	1.87	0.33	1.13	44	14.2	6.34	0.66	0	0
LBNP50	45.02	0	0	3.12	0.54	4.16	0	0.7	30.61	6.22	6.34	0.23	0	0
LBWP51(1)	24.61	0	0	3.58	1.19	5.25	1.19	1.37	40.24	8.94	8.1	0.39	0.47	0
LBWP51(2)	21.35	0	0	5.13	1.45	2.51	0.42	0.75	51.34	0	8.75	0	0	0
LBWP52	25.74	1.22	0	2.27	0.55	1.16	0.23	0.7	47.42	8.64	5.71	0.44	0	0
LBNP53	63.66	3.04	0.11	0.78	1.14	6.06	0	0.54	13.66	8.99	0	0.27	0	0
LBWP54	22.28	0	0.32	3.19	1.28	4.15	0	0.83	55.46	7.04	4.24	0.88	0	0
LBNP55	49.01	2.71	0	1.13	0	1.45	0	0.86	32.92	7.95	0.35	0	0	0
LBWP56	29.15	0	0	1.21	0	1.85	0	0.79	41.25	15.45	2.54	0	0	0
LBWP57	25.42	0	0	1.31	0	2.18	0	1.16	45.1	17.3	0	0	0	0
LBWP58	30.39	0	0.34	3.49	0.79	3.52	0.86	3.13	39.35	7.66	8.99	0.24	0	0

Note: HPWP01(1) means that the part 1 of weathering point 01 of high potassium, HPNP01-(1) means that the part 1 of non-weathering point 01 of high potassium, LBWP01(1) means that the part 1 of weathering point 01 of lead barium, LBNP01(1) means that the part 1 of non-weathering point 01 of lead barium, and LBSWP01(1) means that the part 1 of severe weathering point 01 of high potassium.

References

1. Bzdok, D.; Altman, N.; Krzywinski, M. Points of Significance Statistics versus machine learning. *Nat. Methods* **2018**, *15*, 232–233. [CrossRef]
2. Guo, Y.; Zhan, W.; Li, W. Application of Support Vector Machine Algorithm Incorporating Slime Mould Algorithm Strategy in Ancient Glass Classification. *Appl. Sci.* **2023**, *13*, 3718. [CrossRef]
3. Li, F.; Li, Q.; Gan, F.; Zhang, B.; Cheng, H. Chemical Composition Analysis for Some Ancient Chinese Glasses by Proton Induced X-ray Emission Technique. *J. Chin. Ceram. Soc.* **2005**, *33*, 581–586.
4. Chul, L.; Myungzoon, C.; Seungwon, K.; Kang, H.T.; Du, L.J. Classification of Korean Ancient Glass Pieces by Pattern Recognition Method. *J. Korean Chem. Soc.* **1992**, *36*, 113–124.
5. El-Taher, A. Elemental content of feldspar from Eastern Desert, Egypt, determined by INAA and XRF. *Appl. Radiat. Isot.* **2010**, *68*, 1185–1188. [CrossRef]
6. Won-in, K.; Thongkam, Y.; Pongkrapan, S.; Intarasiri, S.; Thongleurm, C.; Kamwanna, T.; Leelawathanasuk, T.; Dararutana, P. Raman spectroscopic study on archaeological glasses in Thailand: Ancient Thai Glass. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2011**, *83*, 231–235. [CrossRef] [PubMed]
7. Schibille, N.; Gratuze, B.; Ollivier, E.; Blondeau, E. Chronology of early Islamic glass compositions from Egypt. *J. Archaeol. Sci.* **2019**, *104*, 10–18. [CrossRef]
8. Fiorucci, M.; Khoroshiltseva, M.; Pontil, M.; Traviglia, A.; Del Bue, A.; James, S. Machine Learning for Cultural Heritage: A Survey. *Pattern Recognit. Lett.* **2020**, *133*, 102–108. [CrossRef]
9. Wei, J.; Chu, X.; Sun, X.Y.; Xu, K.; Deng, H.X.; Chen, J.G.; Wei, Z.M.; Lei, M. Machine learning in materials science. *Infomat* **2019**, *1*, 338–358. [CrossRef]
10. Zhou, Y.; Hu, Y.; Tao, Y.; Sun, J.; Cui, Y.; Wang, K.; Hu, D. Study on the microstructure of the multilayer glaze of the 16th–17th century export blue-and-white porcelain excavated from Nan'ao-I Shipwreck. *Ceram. Int.* **2016**, *42*, 17456–17465. [CrossRef]
11. Han, M.S. Characteristic Analysis of Chemical Compositions for Ancient Glasses Excavated from the Sarira Hole of Mireuksaji Stone Pagoda, Iksan. *J. Conserv. Sci.* **2017**, *33*, 215–223. [CrossRef]
12. Lin, Y.; Liu, T.; Toumazou, M.K.; Counts, D.B.; Kakoulli, I. Chemical analyses and production technology of archaeological glass from Athienou-Malloura, Cyprus. *J. Archaeol. Sci. Rep.* **2019**, *23*, 700–713. [CrossRef]
13. Oikonomou, A.; Triantafyllidis, P. An archaeometric study of Archaic glass from Rhodes, Greece: Technological and provenance issues. *J. Archaeol. Sci. Rep.* **2018**, *22*, 493–505. [CrossRef]

14. The Official Website of 2022 China Undergraduate Mathematical Contest in Modeling. Available online: http://www.mcm.edu.cn/html_cn/node/5267fe3e6a512bec793d71f2b2061497.html (accessed on 14 May 2023).
15. Kouadri, S.; Pande, C.B.; Panneerselvam, B.; Moharir, K.N.; Elbeltagi, A. Prediction of irrigation groundwater quality parameters using ANN, LSTM, and MLR models. *Environ. Sci. Pollut. Res.* **2022**, *29*, 21067–21097. [CrossRef]
16. Gomah, M.E.; Li, G.; Khan, N.M.; Sun, C.; Xu, J.; Omar, A.A.; Mousa, B.G.; Abdelhamid, M.M.A.; Zaki, M.M. Prediction of Strength Parameters of Thermally Treated Egyptian Granodiorite Using Multivariate Statistics and Machine Learning Techniques. *Mathematics* **2022**, *10*, 4523. [CrossRef]
17. Leonardi, B.; Ajjarapu, V. Development of multilinear regression models for online voltage stability margin estimation. *IEEE Trans. Power Syst.* **2010**, *26*, 374–383. [CrossRef]
18. Tihonov, A.N. Solution of incorrectly formulated problems and the regularization method. *Sov. Math. Dokl.* **1963**, *5*, 1035–1038.
19. Hui, Z.; Hastie, T. Regularization and variable selection via the elastic net. *J. R. Stat. Soc.* **2005**, *67*, 768.
20. Ghosh, S. On the grouped selection and model complexity of the adaptive elastic net. *Stat. Comput.* **2011**, *21*, 451–462. [CrossRef]
21. Li, J.T.; Jia, Y.M.; Zhao, Z.H. Partly adaptive elastic net and its application to microarray classification. *Neural Comput. Appl.* **2013**, *22*, 1193–1200. [CrossRef]
22. Algamal, Z.Y.; Lee, M.H. Applying penalized binary logistic regression with correlation based elastic net for variables selection. *J. Mod. Appl. Stat. Methods* **2015**, *14*, 15. [CrossRef]
23. Hui, Z. The Adaptive Lasso and Its Oracle Properties. *J. Am. Stat. Assoc.* **2006**, *101*, 1418–1429.
24. Van, D.; Lien, T.G.; Verlaat, W.; Wieringen, W.V.; Wilting, S.M. Better prediction by use of co-data: Adaptive group-regularized ridge regression. *Stat. Med.* **2016**, *35*, 368–381.
25. Zhang, F. Combination Model of Enterprise Credit Evaluation Based on XGBoost and Logistic Regression and Its Application. Master's Thesis, Hebei University of Engineering, Handan, China, 2021. (In Chinese) [CrossRef]
26. Jiang, S.; Dai, J. An Improved Elastic Net Estimate for Logistic Regression Models. *Math. Theory Appl.* **2022**, *42*, 108–119. (In Chinese)
27. Anbari, M.E.; Mkhadri, A. Penalized regression combining the L 1 norm and a correlation based penalty. *Sankhya B* **2014**, *76*, 82–102. [CrossRef]
28. Wang, Q.; Li, S.; Li, R.; Ma, M. Forecasting US shale gas monthly production using a hybrid ARIMA and metabolic nonlinear grey model. *Energy* **2018**, *160*, 378–387. [CrossRef]
29. Singh, S.; Mohapatra, A. Repeated wavelet transform based ARIMA model for very short-term wind speed forecasting. *Renew. Energy* **2019**, *136*, 758–768.
30. Wang, C.-C.; Chien, C.-H.; Trappey, A.J. On the application of ARIMA and LSTM to predict order demand based on short lead time and on-time delivery requirements. *Processes* **2021**, *9*, 1157. [CrossRef]
31. Fan, D.; Sun, H.; Yao, J.; Zhang, K.; Yan, X.; Sun, Z. Well production forecasting based on ARIMA-LSTM model considering manual operations. *Energy* **2021**, *220*, 119708. [CrossRef]
32. Li, C.; Fang, X.; Yan, Z.; Huang, Y.; Liang, M. Research on Gas Concentration Prediction Based on the ARIMA-LSTM Combination Model. *Processes* **2023**, *11*, 174. [CrossRef]
33. Jiang, S.; Yang, C.; Guo, J.; Ding, Z. ARIMA forecasting of China's coal consumption, price and investment by 2030. *Energy Sources Part B Econ. Plan. Policy* **2018**, *13*, 190–195. [CrossRef]
34. Ediger, V.Ş.; Akar, S. ARIMA forecasting of primary energy demand by fuel in Turkey. *Energy Policy* **2007**, *35*, 1701–1708. [CrossRef]
35. Dey, B.; Roy, B.; Datta, S.; Ustun, T.S. Forecasting ethanol demand in India to meet future blending targets: A comparison of ARIMA and various regression models. *Energy Rep.* **2023**, *9*, 411–418. [CrossRef]
36. De Gooijer, J.G. Partial sums of lagged cross-products of AR residuals and a test for white noise. *Test* **2008**, *17*, 567–584. [CrossRef]
37. Wang, Y.; Liu, Q. Comparison of Akaike information criterion (AIC) and Bayesian information criterion (BIC) in selection of stock–recruitment relationships. *Fish. Res.* **2006**, *77*, 220–225. [CrossRef]
38. Man, K. Maximum likelihood estimation for a nearly random walk model. *Commun. Stat. Theory Methods* **2000**, *29*, 677–697. [CrossRef]
39. Qureshi, S.A.; Hsiao, W.W.-W.; Hussain, L.; Aman, H.; Le, T.-N.; Rafique, M. Recent Development of Fluorescent Nanodiamonds for Optical Biosensing and Disease Diagnosis. *Biosensors* **2022**, *12*, 1181. [CrossRef]
40. Zheng, C.; Deng, J.; Hong, Z.; Wang, G. Prediction model of suspension density in the dense medium separation system based on LSTM. *Processes* **2020**, *8*, 976. [CrossRef]
41. Lyu, P.; Chen, N.; Mao, S.; Li, M. LSTM based encoder-decoder for short-term predictions of gas concentration using multi-sensor fusion. *Process Saf. Environ. Prot.* **2020**, *137*, 93–105. [CrossRef]
42. Al-Hajj, R.; Assi, A.; Fouad, M. Short-term prediction of global solar radiation energy using weather data and machine learning ensembles: A comparative study. *J. Sol. Energy Eng.* **2021**, *143*, 051003. [CrossRef]
43. Zhu, X.; Li, L.; Liu, J.; Li, Z.; Peng, H.; Niu, X. Image captioning with triple-attention and stack parallel LSTM. *Neurocomputing* **2018**, *319*, 55–65. [CrossRef]
44. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]
45. Olah, C. Understanding lstm Networks. 2015. Available online: https://colah.github.io/posts/2015-08-Understanding-LSTMs/ (accessed on 24 March 2023).

46. Wu, X.; Zhou, J.; Yu, H.; Liu, D.; Xie, K.; Chen, Y.; Hu, J.; Sun, H.; Xing, F. The development of a hybrid wavelet-ARIMA-LSTM model for precipitation amounts and drought analysis. *Atmosphere* **2021**, *12*, 74. [CrossRef]
47. Xu, D.; Zhang, Q.; Ding, Y.; Zhang, D. Application of a hybrid ARIMA-LSTM model based on the SPEI for drought forecasting. *Environ. Sci. Pollut. Res.* **2022**, *29*, 4128–4144. [CrossRef] [PubMed]
48. Wei, X.; Shahani, N.M.; Zheng, X. Predictive Modeling of the Uniaxial Compressive Strength of Rocks Using an Artificial Neural Network Approach. *Mathematics* **2023**, *11*, 1650. [CrossRef]
49. Xu, P. Prediction of Per Capita Ecological Carrying Capacity Based on ARIMA-LSTM in Tourism Ecological Footprint Big Data. *Sci. Program.* **2022**, *2022*, 6012998. [CrossRef]
50. Manowska, A.; Rybak, A.; Dylong, A.; Pielot, J. Forecasting of Natural Gas Consumption in Poland Based on ARIMA-LSTM Hybrid Model. *Energies* **2021**, *14*, 8597. [CrossRef]
51. Huang, Y.; Fan, J.; Yan, Z.; Li, S.; Wang, Y. Research on early warning for gas risks at a working face based on association rule mining. *Energies* **2021**, *14*, 6889. [CrossRef]
52. Bukhari, A.H.; Raja, M.A.Z.; Sulaiman, M.; Islam, S.; Shoaib, M.; Kumam, P. Fractional neuro-sequential ARFIMA-LSTM for financial market forecasting. *IEEE Access* **2020**, *8*, 71326–71338. [CrossRef]
53. Salmerón, R.; García, C.; García, J. Variance inflation factor and condition number in multiple linear regression. *J. Stat. Comput. Simul.* **2018**, *88*, 2365–2384. [CrossRef]
54. Burnham, K.P.; Anderson, D.R. Multimodel inference: Understanding AIC and BIC in model selection. *Sociol. Methods Res.* **2004**, *33*, 261–304. [CrossRef]
55. Hassani, H.; Yeganegi, M.R. Sum of squared ACF and the Ljung–Box statistics. *Phys. A Stat. Mech. Appl.* **2019**, *520*, 81–86. [CrossRef]
56. Lee, T. Wild bootstrap Ljung–Box test for cross correlations of multivariate time series. *Econ. Lett.* **2016**, *147*, 59–62. [CrossRef]
57. Hollas, B. An analysis of the autocorrelation descriptor for molecules. *J. Math. Chem.* **2003**, *33*, 91–101. [CrossRef]
58. Angel, E.; Zissimopoulos, V. Autocorrelation coefficient for the graph bipartitioning problem. *Theor. Comput. Sci.* **1998**, *191*, 229–243. [CrossRef]
59. Goudet, J. FSTAT (version 1.2): A computer program to calculate F-statistics. *J. Hered.* **1995**, *86*, 485–486. [CrossRef]
60. Weir, B.S.; Hill, W.G. Estimating F-statistics. *Annu. Rev. Genet.* **2002**, *36*, 721–750. [CrossRef]
61. Weaver, B.; Wuensch, K.L. SPSS and SAS programs for comparing Pearson correlations and OLS regression coefficients. *Behav. Res. Methods* **2013**, *45*, 880–895. [CrossRef]
62. Halunga, A.G.; Orme, C.D.; Yamagata, T. A heteroskedasticity robust Breusch–Pagan test for Contemporaneous correlation in dynamic panel data models. *J. Econom.* **2017**, *198*, 209–230. [CrossRef]
63. Jeong, J.; Lee, K. Bootstrapped White's test for heteroskedasticity in regression models. *Econ. Lett.* **1999**, *63*, 261–267. [CrossRef]
64. Baum, C.; Cox, N. WHITETST: Stata Module to Perform White's Test for Heteroskedasticity. 2002. Available online: https://econpapers.repec.org/software/bocbocode/s390601.htm (accessed on 25 March 2023).
65. Koklu, M.; Taspinar, Y.S. Determining the Extinguishing Status of Fuel Flames With Sound Wave by Machine Learning Methods. *IEEE Access* **2021**, *9*, 86207–86216. [CrossRef]
66. Su, C.; Wang, J. Research on composition analysis and type identification of ancient glass products based on data mining. *Autom. Mach. Learn.* **2022**, *3*, 63–71. [CrossRef]
67. Sun, C.; Li, Z. Analysis and Identification of the Composition of Ancient Glass Objects based on Statistical Research and Machine Learning Algorithms. *Highlights Sci. Eng. Technol.* **2023**, *39*, 1412–1418. [CrossRef]
68. Pu, Q.; Jiang, L.; Liu, Z.; Wang, X.; Liu, Z. Research on Classification of Ancient Glass Products Based on Machine Learning. In Proceedings of the 2022 International Conference on Information Technology, Communication Ecosystem and Management (ITCEM), Bangkok, Thailand, 19–21 December 2022.
69. Bai, D. Comparative study on chemical composition of ancient glass based on machine learning and deep learning. *Highlights Sci. Eng. Technol.* **2022**, *22*, 234–240. [CrossRef]